The Alberta Journal of Educational Research

Vol. X, No. 3 *September, 1964*

THE COMMITTEE ON EDUCATIONAL RESEARCH
Faculty of Education
University of Alberta

VALIDITY OF REGRESSION EQUATIONS AFTER SIX YEARS TO PREDICT FRESHMAN SUCCESS IN ENGINEERING

By

D. B. Black

In earlier articles by the author (1960, 1961), regression equations were developed to predict freshman success in Engineering at the University of Alberta, and these equations were validated on a class of Engineering freshman a year later. The sample used to develop the equations was the 1956 entering class while the validating class was the 1957 entering class. This study seeks to test the effectiveness of these same equations after six years for the 1962 entering class of freshman Engineers, to develop new equations to meet the changing freshman curriculum, and because of the increasing problem of early university admission and the trend towards school accreditation, to examine the effectiveness of the Principal's Ratings to predict freshman Engineering success.

The sample used in this study was 102 freshman engineers who had graduated from Alberta high schools in the spring of 1962, who entered the Faculty of Engineering that fall, and for whom a full record of the seven prescribed Grade XII Departmental examination scores (English 30, Social Studies 30, Mathematics 30 and 31, Chemistry 30, Physics 30 and a foreign language 30), were available. Similarily, only those entering freshman for whom the Department of Education records indicated Principal's Confidential Ratings for the equivalent subjects used for admission to the Faculty were used in the study.

Several changes have taken place since the 1956 study. First, the Departmental scores in Mathematics 31 are now based on an examination rather than the teachers' mark or Principal's Rating. In 1957, the Faculty of Engineering replaced the 1956 Drawing 2 and 4 courses with a new course Drawing 2. Courses in Chemistry, Physics and Mathematics have been retained. Drawing 2 became Drafting and Graphics 151 while Engineering Mechanics 1 evolved into two courses, Statics (E.M. 150) and Particle Dynamics (E.M. 151). The two survey courses, C.E. 5 and 6 have been reduced to a single laboratory (field) course, C.E. 153, which was not included in the current study. The half course, Eng. P. 159, Orientation to Engineering Profession, has not been included in the current study as well.

TABLE I

COMPARISON OF 1957-58 AND 1962-63 GRADE XII PREDICTOR VARIABLE SAMPLE MEANS TO 1956-57 SAMPLE MEANS AND 1962-63 PRINCIPAL'S RATINGS TO ACHIEVED GRADES

Grade XII Predictor	1956-57 Sample N=131		1957-58 Sample N=245		1962-63 Sample (N=102)				Significance of Differences of Means			
					Ach'd Grades		Pr's Rating		1956 and 57 Samples	1956 and 62 Samples		
										Ach'd Grades	Pr's Rating	1962 Principal's Rating and Achieved Grades
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD				
English 30	65.2	10.2	63.8	9.6	67.0	9.5	64.8	10.8	NSD	NSD	NSD	NSD
Social Studies 30	73.3	10.5	69.9	10.9	70.9	9.8	71.1	10.4	1%	NSD	NSD	NSD
Mathematics 30	74.4	10.3	72.6	9.2	74.0	10.6	73.9	9.5	NSD	NSD	NSD	NSD
Mathematics 31	69.4	14.1	70.3	11.3	68.6	11.7	68.5	10.6	NSD	NSD	NSD	NSD
Physics 30	72.8	10.6	70.6	9.4	73.8	10.1	69.3	10.5	NSD	NSD	5%	5%
Chemistry 30	73.2	11.9	72.1	9.4	74.3	9.8	70.8	10.5	NSD	NSD	NSD	5%
Foreign Language 30	66.2	11.6	65.7	10.1	67.2	10.7	63.3	12.3	NSD	NSD	NSD	5%
Mathematics Average	71.6	10.8	71.3	8.7	71.6	10.1	71.5	9.1	NSD	NSD	NSD	NSD
Science Average	72.7	10.4	71.1	8.5	74.0	8.9	69.9	9.1	NSD	NSD	5%	5%

The method of the study was to use the Grade XII Departmental scores, both those actually achieved on the Departmental Examinations (Act.), and the Principal's Ratings (P.R.), in the 1956 regression equations to predict the freshman success of the 1962 sample of Engineering freshman. Where the Engineering courses or criterion variables were approximately the same as in 1956, these predicted and achieved grades in these courses were correlated. This was done on the assumption that the grading practices of the Department of Education and the Faculty of Engineering, respectively, have not changed since 1956. This assumption will be tested. The second assumption is that notwithstanding changes in the content both in Grade XII and freshman courses in six years, the skills and abilities necessary to complete a course in Mathematics, for example, will have remained unchanged, and that the scores reported from examinations are a good representation of this ability.

The second part of the study will deal with the development of new regression equations to meet the current curriculum in freshman Engineering. The Dominion Bureau of Statistics Stepwise Regression Computer Program for an IBM 1620 computer was used to develop these equations. All calculations were done at the University of Alberta's Computer Center.

Findings:

To test the assumption that the grading practices of the Department and the University had not changed, the means, standard deviations were calculated for the 1962 sample for all predictor and criterion variables including the Principal's Ratings. Statistics for the 1956 and 1957 samples are those as reported earlier by the author (Black, 1961). These data and analyses are reported in Tables I and II.

Table I shows that with one exception there are no statistically significant differences between the Departmental achieved grades of the 1956 and 1962 freshman samples. Significant differences at the 5% level were found between 1962 achieved and Principal's Ratings Departmental scores in Physics 30, Chemistry 30 and Foreign Language 30, and because the significant differences in both Physics and Chemistry grades, in Science average as well. It is interesting to note that in every instance, but Social Studies 30, the mean reported Principal's Rating for the 1962 sample was lower than the actually achieved mean. In the case of Social Studies, the difference was two tenths of a mark. This consistent trend to report lower grades suggests strongly (at 1% level of confidence using the Sign Test) that average high school teacher's ratings for Engineering freshman, at least, were consistently lower than those

actually achieved. This fact plus the statistically significant differences for primary science predictors, and Physics 30 in particular, should be of great importance to the Faculty of Engineering in the event that it considers early admission based on Principal's Ratings. In terms of the conventional t-Test, there are no significant differences in overall mean high school performance of the 1956 and 1962 freshman classes. It is concluded then, that the assumption of a consistent grade admission pattern in Grade XII is a valid one.

The assumption of a consistent grading pattern in University freshman engineering grades would not appear to be as valid. These data are reported in Table II. In 1957, there appeared to be a change downward in Physics, Chemistry and Drafting marks, while mean Mathematics marks climbed. The 1962 marks tended to follow the 1957 pattern with the exception of Mathematics which dropped below the 1956 mean and which was consistent with decline in 1957 mean marks in Physics and Chemistry. Only mean marks in Engineering Mechanics remained consistent. The introduction of the English course as indicated in the 1962 sample with a mean grade of 51.6 is reflected in the consistent downward trend in Engineering Average. The Average was not significantly different between the 1956 and 1957 classes, but it was highly significant for both the 1956 and 1957 and the 1962 mean freshman performance. It can be concluded that inasmuch as there were no significant differences in the capacity of the 1962 freshman class and those entering in 1956 and 1957 as indicated by the Grade XII marks, the marking practices of the University has changed and this change would appear to have been effected primarily in the Arts and Sciences courses beginning in 1957.

Although there has evidently been a change in the level of criterion grades, this, in turn, could be viewed as a consistent change. Therefore, in correlating the predicted grades for 1957 and 1962 freshman classes, the effectiveness of the equations to predict freshman success may not be invalidated because correlation coefficients are dependent on relative ranking in a distribution rather than absolute value of the scores. The validity coefficients of the 1956 equations to predict 1957 and 1962 freshman grades are reported in Table III. Two sets of equations are used, seven variable equations using each of the seven Grade XII Departmental prerequisites for admission, and a five variable equation which uses Mathematics Average ($\overline{X}$) instead of Mathematics 30 and 31 scores and Science Average ($\overline{X}$) instead of Chemistry 30 and Physics 30 scores. The 1962 freshman grades are predicted using both actual Grade XII marks and the Principal's Ratings.

TABLE II

COMPARISON OF 1956 FRESHMAN SAMPLES WITH 1957 AND 1962 FRESHMAN SAMPLES: FRESHMAN MEANS GRADES

Engineering Freshman Variable	Course Number			Freshman Sample						Significance of Difference of Mean		
				1956 N=131		1957 N=245		1962 N=102				
	'56	'57	'62	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	1956 & '57 Sample	1956 & '62 Sample	Sample 1957 & '62
Engineering Average	...	...	...	63.3	10.8	61.4	11.3	57.5	13.1	NSD	1%	1%
Mathematics	2	2	202	60.5	17.7	68.0	15.9	53.1	18.5	1%	1%	1%
Physics	21	21	230	61.6	20.2	54.6	16.5	57.6	18.8	1%	NSD	NSD
Chemistry	40	40	232	60.4	14.1	57.1	13.5	55.5	12.6	5%	1%	NSD
Drafting and Graphics	2*	2	151	67.2	8.8	62.1	10.8	61.3	9.7	1%	1%	NSD
	4*	2	151	69.8	11.2	62.1	10.8	61.3	9.7	1%	1%	NSD
Engineering Mechanics	1	1	150	62.2	15.9	62.5	15.1	59.7	18.4	NSD	NSD	NSD
	1	1	151	62.2	15.9	62.5	15.1	61.6	23.8	NSD	NSD	NSD
Civil Engineering	5	5	...	62.5	13.0	60.9	13.5	...	...	NSD	...	...
	6	6	...	64.6	8.2	65.8	6.8	...	...	NSD	...	...
English	...	...	250	...	...	...	...	51.6	8.5	...	...	...

*NOTE: In 1957 Drawing 4 was dropped as a course. Comparisons are made to the single Drawing Course in 1957 ,1962.

TABLE III: ACTUAL AND ESTIMATED VALIDITY COEFFICIENTS AND STANDARD ERRORS OF ESTIMATE FOR PREDICTED AND ACHIEVED FRESHMAN ENGINEERING COURSES OF ENTERING CLASSES IN 1956, 1957 AND 1962 BASED ON GRADE XII DEPARTMENTAL GRADES AND PRINCIPAL'S RATINGS

	Criterion Calendar Course Number			1956 Sample (N=131)				Obtained Validity Coefficients for 1957 Sample (N=245)				Obtained Validity Coefficients for 1962 Sample (N=102) Predicted from 1956 Prediction Equations								Validity Coefficients* for New Equations based on 1962 Class for 1962 Freshman Courses							
				Actual Marks				Actual Marks				Actual Marks				Principal's Ratings				Actual Marks				Principal's Ratings			
				7 Variable		5 Variable		7 Variable		5 Variable		7 Variable		5 Variable		7 Variable		5 Variable		7 Variable		5 Variable		7 Variable		5 Variable	
				R.	SE.est.	R.	SE.est.	R.	SE.est.	R.	SE.est.	R.	SE.est.	R.	SE.est.	R.	SE.est.	R.	SE.est.	R.	SE.est.	R.	SE.est.	R.	SE.est.	R.	SE.est.
Year	1956	1957	1962																								
Engineering Average	-	-	-	.69	7.8	.67	8.0	.56	8.6	.55	8.3	.73	6.8	.72	6.9	.64	7.9	.64	7.9	.76	8.8	.75	8.8	.67	10.1	.66	10.1
Mathematics	2	2	202	.69	12.8	.68	13.0	.57	15.1	.56	14.7	.66	10.5	.65	10.9	.61	11.4	.60	11.6	.69	14.0	.68	14.0	.63	14.9	.62	14.9
Physics	21	21	230	.65	15.4	.63	15.7	.61	14.5	.64	14.3	.63	11.4	.66	10.8	.51	13.1	.56	11.4	.70	13.9	.70	13.9	.58	15.8	.57	15.8
Chemistry	40	40	232	.70	10.1	.69	10.2	.59	11.6	.61	11.4	.73	6.5	.74	6.4	.61	7.8	.60	6.9	.76	8.5	.76	8.4	.64	10.1	.62	10.2
Drafting and Graphics	2	2	151	.51	7.6	.48	7.7	.41	9.8	.36	11.2	.61	6.0	.58	6.3	.52	6.7	.48	7.0	.67	7.5	.62	7.8	.54	8.5	.51	8.6
	4	2	151	.41	10.1	.42	10.2	.42	12.4	.40	12.5	.59	6.2	.51	6.8	.49	6.9	.45	7.2	-	-	-	-	-	-	-	-
Engineering Mechanics	1	1	150	.54	13.4	.52	13.6	.51	12.7	.46	13.5	.49	12.0	.48	12.2	.47	12.4	.44	12.7	.61	15.0	.58	15.4	.55	15.9	.48	16.5
	1	1	151	.54	13.4	.52	13.6	.51	12.7	.46	13.5	.66	12.8	.58	15.4	.56	15.5	.53	16.4	.71	17.3	.70	17.6	.62	19.4	.60	19.6
Civil Engineering	5	5	-	.57	10.7	.55	10.9	.45	12.2	.39	12.9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	6	6	-	.12	8.1	.08	8.2	.12	7.6	.11	6.8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
English	-	-	250	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.60	7.1	.59	7.0	.54	7.2	.55	7.5

*Coefficients are based on maximum R achieved by using all Grade XII predictor variables in regression equation.

The validity coefficients for the 1956 and 1957 samples have been reported earlier (Black, 1961) and will be used to compare the validity coefficients obtained for the 1962 sample. It will be noted that in the 1956 validation study using actual marks for the seven variable predictor equations, all but one of the obtained validity coefficients were lower than those for the 1957 sample. In contrast, for the 1962 sample, five of the eight validity coefficients were numerically higher when actual marks were used to predict success. The predictions using the 1962 Principal's Ratings showed that only three of eight validity coefficients were numerically higher. In every instance but two for both actual and Principal's Rating predictors, the validity coefficients obtained for the 1962 sample were numerically higher than those obtained for the 1957 sample.

The same general pattern held true for the five variable predictions. Of the validity coefficients for the 1962 sample six of eight are numerically greater than the 1956 sample; for the 1957 sample, seven of eight are equal to or higher numerically. Of the Principal's Ratings for the five variable predictions three of eight are equal to or higher numerically than the actual mark validity coefficients for the 1956 sample, and six of eight are higher than the 1957 sample.

In no instance for any comparison were differences found that were statistically significant at even the 5% level. However, the general pattern of differences favoring the 1962 sample would confirm the general observation that the regression equations developed from the 1956 sample worked equally well for the 1962 sample. The evidence does suggest a stability that was much greater than suspected and would thus favor operational implementation of the regression equations within the admission procedures of the University.

By the same token, the Principal's Ratings, as predictor variables, performed as would have been anticipated from earlier studies and further support the case for use of these marks for early provisional admission procedures.

The final part of Table III reports the validity coefficients and standard errors of measurement for new regression equations for the current freshman offering of the Faculty of Engineering. In every instance, when such comparisons can be made, the actual mark of seven and five predictor validity coefficients are equal to or numerically higher than those obtained for the 1956 sample. The validity coefficients for the equations using Principal's Ratings are slightly lower.

Table IV

RELATIVE BETA WEIGHTS* AND MULTIPLE CORRELATION COEFFICIENTS FOR SEVEN GRADE XII DEPARTMENTAL SCORES TO PREDICT SEVEN FRESHMAN ENGINEERING SCORES

Engineering Criterion	Grade XII Departmental Score Predictor							R	Standard Error of Estimate
	English 30	Soc. St. 30	Math. 30	Math. 31	Chem. 30	Physics 30	For. Lang. 30		
Engineering Average	—31	39	15	9	33	75	4	.762	8.8
Mathematics 202	—54	59	59	24	43	113	17	.687	14.0
Physics 230	—73	62	—11	26	101	142	14	.701	13.9
Chemistry 232	—10	32	14	4	66	47	—14	.760	8.5
Drafting and Graphics 151	—24	32	—3	7	—4	55	4	.666	7.5
Engineering Mechanics 150	—93	67	74	—14	14	138	—18	.614	15.0
Engineering Mechanics 151	—124	64	33	41	69	273	42	.713	17.3
English 250	23	25	—2	3	—1	8	2	.597	7.1

*Multiplied by 100 to remove decimals.

As evidenced earlier by changes in the freshman Engineering curriculum and in the marking patterns of the University courses since 1956, there is need for new regression equations to replace those developed earlier. This has been done using the 1962 freshman sample of 102 cases. The computer program used in this study was such that each regression equation would be examined as each successive predictor variable in order of contribution to the prediction was added to the equation. If the program ran to its full length, i.e., all variables were used, a regression equation with a maximum multiple correlation coefficient was obtained with the criterion in which each of the variables was weighted to give an optimal relationship. It is evident from Table III, that the actual grades rather than Principal's Ratings on all seven required Grade XII Departmental subjects, afford the best predictors of freshman success at the University. Using all seven Grade XII Departmental scores, the relative weightings (Beta weights) of the different subjects for the eight 1962 freshman Engineering criteria were calculated and are reported in Table IV.

It will be noted that Physics 30 is the highest and most frequent positive contributor to the prediction of freshman success followed closely by Social Studies 30 and Chemistry 30. The two Mathematics grades do not contribute as significantly as might be anticipated. The foreign language 30 grade makes only a small contribution. Contrasting is the consistent inverse relationship between English 30 grades and all freshman grades, except English 250. Care should be taken in interpreting these data. They show only the relationships of the various predictors to the prediction of success. They do not relate in any way to the necessity of various courses as prerequisites for study. Therefore, while it is readily apparent that there is an inverse relationship between success in high school English and Engineering success, this is no way to deny the importance of English even to Engineers. The data do suggest, however, that whatever skills are being measured and presumably have been taught in Grade XII English as measured by the Departmental examinations are not those required by prospective and successful Engineering freshmen. Conversely, one is puzzled by the consistent positive contribution of Social Studies 30. It is possible that the communicative skills demanded of Engineering students are best taught and measured in this course. When one considers that even in English 250, the Social Studies 30 mark contributes about equally as the English 30 mark to the predicton of this course, the above hypothesis would seem to have some merit. What the data in Table IV do suggest is that in the selection of Engineering freshmen, the Faculty of Engineering could do well to

Table V

OPERATIONAL REGRESSION EQUATIONS TO PREDICT FRESHMAN ENGINEERING GRADES USING SEVEN GRADE XII DEPARTMENTAL EXAMINATION SCORES AS PREDICTOR VARIABLES

Engineering Criterion	Type of Score	Regression Equations: Grade XII Predictors: English 30	Soc. St. 30	Math. 30	Math. 31	Chem. 30	Physics 30	For. Lang. 30	Constant	Standard Error of Estimate	R^2
Engineering Average	Act.	—.23	.32	.....	.....	.29	.70	.....	—22.4	8.8	.75
	P.R.	.....	.19	.26	.28	.....	.37	.....	—19.0	10.1	.66
Mathematics 202	Act.	.....	.30	.45	.....	.....	.74	.....	—55.7	14.1	.67
	P.R.	.....	.....	.58	.41	.....	.33	.....	—40.1	14.9	.61
Physics	Act.	—.38	.35	.....	.....	.55	.85	.....	—45.4	13.9	.70
	P.R.	.....	.....	.....	.40	.34	.50	.....	—28.5	15.7	.57
Chemistry 232	Act.	.....	.21	.....	.....	.50	.45	.....	—30.0	8.5	.75
	P.R.	.....	.....	.....	.31	.53	.....	.....	—31.9	10.2	.60
Drafting and Graphic 151	Act.	—.21	.32	.....	.....	.....	.52	.....	14.0	7.5	.66
	P.R.	.....	.19	.....	.....	.....	.39	.....	20.5	8.4	.53
Engineering Mechanics 150 (Statics)	Act.	—.47	.35	.37	.....	.....	.75	.....	—16.2	15.0	.61
	P.R.	.....	.....	.....	.41	.....	.60	.....	—10.1	16.0	.51
Engineering Mechanics 151 (Particle Dynamics	Act.	—.35	.....	.....	.34	.41	1.19	.....	—56.3	17.3	.70
	P.R.	.....	.....	.47	.60	.....	.59	.....	—55.0	19.3	.60
English 250	Act.	.25	.29	.....	.....	.....	.11	.....	5.3	7.1	.59
	P.R.	.19	.20	.....	.....	.....	.13	.....	16.2	7.1	.53

ask only for passing grades in both English 30 and a foreign language 30 and exclude these marks from any inclusion in the calculations that determine whether a freshman will or will not be admitted to the Faculty.

It is also apparent from earlier studies that prediction equations based on all predictor variables can be successfully reduced with little loss in predictive efficiency. This has the advantage of making the predictions much easier to calculate in the high schools which do not normally have access to electric calculators or computer services. To reduce the number of predictor variables, the relative beta weight of each predictor added successively to the equation was tested for its significance greater than zero. When this hypothesis (Null hypothesis) could be accepted at the 10% level of significance, subsequent predictor variables by virtue of the programs selection of variables in order of their contribution to the prediction were dropped. These "reduced" regression equations are identified as "operational" regression equations and are reported in Table V ready for use in the schools. Equations have been calculated for both actual marks and Principal's Ratings. The five predictor variable equations are not reported because of their limited use due to the prescriptive nature of the prerequisite Grade XII scores for admission. The seriousness of the reduction from seven to three or four predictor variables is readily apparent by comparing the multiple correlation coefficient and standard errors of estimate in Table IV to those for actual grades in Table V. In no case was the multiple correlation coefficient lowered by more than two points or the standard error of estimate increased by more than two tenths of a score point.

Summary

This study has sought to evaluate the effectiveness of regression equations developed in 1956 to predict success for a freshman engineering class six years later. It was found that they remained effective over this period. Moreover, when Principal's Ratings were used instead of marks actually obtained on the Grade XII Departmental examinations, the predictions remained sufficiently high to be useful for admission procedures. The findings offer promise for their operational incorporation in the admission practices of the Faculty and have implications for inaugurating a practice of early (Easter) provisional admission to the Faculty pending successful completion of the June Departmental examinations. Certainly the stability of the equations which was found to be higher than would have been suspected, offers further encouragement in this direction.

New regression equations have been developed to accommodate changes in the curriculum and grading practices of the freshman year. It has been shown that those equations can be modified with little loss in predictive efficiency to make them more practical for use by school counsellors.

BIBLIOGRAPHY

Black, D. B. "The Prediction of Freshman Success in the University of Alberta from Grade XII Departmental Results." *Alta. Journal of Ed. Res.*, VI:1 (March, 1960), 30-53.

"Differential Grade Prediction: A Note on Its Effectiveness for a Freshman Class of Engineers." *Alta. Journal of Ed. Res.*, VII:2 (June, 1961), 86-92.

AN INVESTIGATION INTO THE KNOWLEDGE OF MATHEMATICS OF ALBERTA STUDENTS, BASED ON THE GRADE NINE DEPARTMENTAL EXAMINATION WRITTEN IN JUNE, 1960

Ellen A. Laws

SCORES: GRADE IX EXAMINATIONS

Each year in June the grade IX high school entrance examinations are conducted under the direction of the High School Entrance Examinations Board. Students write papers in language, literature, social studies, science, mathematics, and reading comprehension. They pass or fail on the basis of their achievement on these papers. In addition they write a test of general aptitude.

One of the major problems of the Examinations Board is that of setting standards for judging satisfactory achievement. The present practice of scaling final marks has received much criticism on the basis that standards have fallen in recent years. The writer is not prepared to judge the accuracy of this contention, but it must be remembered that economic pressure is forcing students to remain in high school who, twenty or thirty years ago would have dropped out at the end of grade VIII and found gainful employment. The result is that students of all but the lowest levels of academic ability are writing grade IX examinations.

The Royal Commission on Education investigated the field of departmental examinations (1). It found that in the years from 1906 to 1935, the marks awarded to a student depended primarily upon the difficulty of the examination. Passing marks based on raw scores, varied from year to year. Various systems of bonus marks were used to achieve a reasonable distribution of marks, but lack of consistency in their application rendered the system unsatisfactory.

The present policy of the Department of Education is as follows:

> The grading of the pupils is based on the assumption that in a LARGE group the students from year to year are about the same in ability and achievement, that is to say, from one year to the next there is the same percentage of average, of dull and of bright pupils. It would follow that the same proportion of pupils should receive high marks each year and that the percentage of failures should not vary. The actual proportion of pupils in each category was set several years ago, and except for occasional minor changes, is used from year to year and applies to almost all of the subjects (2).

The following table shows the distribution of letter gradings which was used for grade IX pupils in 1960 when this study was begun*—

TABLE I

DISTRIBUTION OF LETTER GRADINGS TO GRADE IX PUPILS

Letter gradings given to pupils	Percentage marks used to derive letter gradings	Percentage of total number of pupils
H	80-100	10
A	65- 79	25
B	50- 64	35
C	40- 49	20
D	0- 39	10

*The present distribution became effective in 1962.

When departmental examinations are marked a raw score for each paper is established. This is later converted to a scaled score. The student with the highest raw score for the province is given a scaled score which is usually one hundred percent, while the student with the lowest raw score is assigned a scaled score near zero. The raw scores are arranged in rank order. The top ten percent are equated to percentage gradings ranging from one hundred percent down to eighty percent by means of a linear graph. These students are given an H grading. The same method is used in equating percentage scores and gradings for other categories.

While percentage scores are kept on record at the Department of Education, only letter gradings appear on the diplomas sent to the individual pupils. To the teachers receiving these pupils in grade X, the letter gradings and the percentage marks which they represent indicate only the relative position of the pupil to the entire student body of the province. A B pass mark, for instance, which is shown on the statement issued to the pupils as indicating a score of from fifty to sixty-four percent, does not mean that the raw score was necessarily in the fifty to sixty-four range, or that the pupil knew at least one-half of the material tested on the examination paper. It seems desirable for grade X teachers to have a fairly clear picture of their pupils' ability in relation to the grade IX course, as well as their ability in relation to the other pupils in the province. It was with this object in mind that the writer launched an investigation into the grade IX Mathematics papers written in the year 1960.

SCORES: GRADE IX MATHEMATICS EXAMINATIONS

The data for the study came from a proportional systematic sample of 389 papers, selected with the help of the Examinations Branch from the total of all Grade IX mathematics papers written in the Province of Alberta in June, 1960. The sample contained the same proportion of papers in each of the categories, H, A, B, C, and D as the diplomas issued in the province that year.

The first step in the investigation of the sample papers was the preparation of an item analysis. The principles involved in the solution of each examination question were itemized and these were used as the basis for scoring the sample papers.

The grade IX course in mathematics includes four areas of study, arithmetic, algebra, geometry, and graphic study which is the visual presentation and interpretation of numerical data. A separate frequency distribution was made for each category of student, H, A, B, C, D, in each of the four areas of study. Each item was checked according to whether the principle was correct, incorrect or not attempted. In solutions involving more than one principle, each principle was examined independently of those that preceded or followed it, and credit was given for knowledge of that principle without regard to the values which had been assigned when the paper was received. Where the same principle occurred more than once on the examination paper, the totals were combined.

The performance of each category of student, H, A, B, C, D, was computed separately for each of the four areas of study. The results are summarized in Tables II-VI. To arrive at the figures shown in the tables, for each category of student the total number of correct responses in a single area of study, e.g. arithmetic, was expressed as a percentage of the total possible in the area. For comparison, each table also includes the percentage implied by the letter grading issued to the student.

TABLE II

RELATIVE COMPETENCE OF THE H GROUP OF STUDENTS IN THE FOUR AREAS OF STUDY

H (80 - 100 percent)		
Arithmetic	78.51	per cent
Algebra	86.74	per cent
Geometry	81.36	per cent
Graphic Study	83.97	per cent

TABLE III

RELATIVE COMPETENCE OF THE A GROUP OF STUDENTS IN THE FOUR AREAS OF STUDY

A (65 - 79 percent)		
Arithmetic	60.19	per cent
Algebra	67.75	per cent
Graphic Studies	60.29	per cent
Geometry	77.12	per cent

TABLE IV

RELATIVE COMPETENCE OF THE B GROUP OF STUDENTS IN THE FOUR AREAS OF STUDY

B (50 - 64 percent)		
Arithmetic	43.09	per cent
Algebra	46.33	per cent
Geometry	44.23	per cent
Graphic Study	60.83	per cent

TABLE V

RELATIVE COMPETENCE OF THE C GROUP OF STUDENTS IN THE FOUR AREAS OF STUDY

C (40 - 49 percent)		
Arithmetic	24.99	per cent
Algebra	23.77	per cent
Geometry	30.91	per cent
Graphic Study	42.95	per cent

TABLE VI

RELATIVE COMPETENCE OF THE D GROUP OF STUDENTS IN THE FOUR AREAS OF STUDY

	D (0 - 39 percent)	
Arithmetic	15.04	per cent
Algebra	10.53	per cent
Geometry	17.96	per cent
Graphic Study	29.73	per cent

CONCLUSIONS

It must be made clear that any conclusions drawn in this study are founded entirely on the results of one Departmental examination. Therefore, they must be qualified with the observation that this is what the students knew, or did not know, in 1960. In no way is this investigation intended to establish final conclusions regarding the ability of all students who have written, or will continue to write the grade IX examination.

However, the unsatisfactory standard of work exhibited by the B, C, and D groups in this investigation poses the same question that has been asked by other investigators—"What can be done to raise the level of achievement in mathematics?" Is the answer to be found only in improved instructional procedures, in more intensive drill on mechanical processes, and in stricter discipline in mathematical routines? Or should we take a second look at the mathematics curriculum and ask ourselves whether our present course is designed to do what it is intended to do. Mathematics, to justify its place in the school curriculum, must contribute to the satisfaction of the needs of the student in the fields of personal living, and social and economic relationships.

REFERENCES

1. Report on the Royal Commission on Education. Edmonton: Government of Alberta, The Queen's Printer, 1959.
2. Department of Education, Province of Alberta. "Curriculum News Letter, No. 9." Edmonton: The Queen's Printer, February, 1958.

REFERENCE GROUP BEHAVIOR AMONG HIGH SCHOOL STUDENTS

HENRY ZENTER

University of Alberta, Calgary

This paper examines certain aspects of reference group behavior among high school students of varied racial, socio-economic and ecological description in both Oregon and Alberta.

Although the concept of reference group behavior was recently formulated by Merton, the idea that an individual is influenced in his behavior by other persons and groups that are in some sense meaningful to him has long been held as a fundamental principle of behavior among sociologists and social psychologists.[1] George H. Mead, in particular, attributed central importance to 'significant others' in the socialization of the child.[2] Thus, while the underlying notion of influence upon behavior emanating from others has led to extensive speculation and considerable research in sociology in recent years, it would, no doubt, readily be agreed that we are far from an adequate understanding of the circumstances under which given individuals will take given categories of 'significant others' into account. Neither is it known to what extent 'significant others' influence the behavior of given individuals under given circumstances. Perhaps least of all among high school students.

The present inquiry, therefore, is focussed upon the relationship between the high school students and a number of selected 'significant other' reference categories respecting two major role-taking decisions facing students.[3] The two role-taking situations selected for study include: (1) graduating from high school; and (2) going on to further training after high school has been completed. The data reported here are taken from a series of more

1. Robert K. Merton, *Essays in Social Theory and Social Structure*, rev. ed., New York. The Free Press of Glencoe, Inc., p. 299 ff.
2. George H. Mead, *Mind, Self and Society*, University of Chicago Press, 1934.
3. Ralph H. Turner, "Role-Taking, Role Standpoint, and Reference-Group Behavior", *American Journal of Sociology*, LXI (1956), 316-328. Turner identifies three major types of reference groups that lie on a continuum. Following Sherif and Shibutani, he places the *identification group* at one end of the continuum since it is seen as the source of values with the result that the individual adopts the member's standpoint as his own. At the opposite end of the continuum is a second type called the *interaction group*, so designated since the members of such a group constitute merely conditions to the behavior of the actor. In between is a third type of group called the *valuation group*. The latter is differentiated from the remaining two types in that its effect upon the individual's behavior will be determined by the valuation which his more basic orientation leads him to place upon it. It is the latter type which would appear to subsume the reference categories employed in the present study. Included are: relatives, teachers, school friends, friends of the family, other people liked, and neighbors. Since there is no theoretical justification for the assumption that these classes of persons constitute real groups, the term "reference category" has been employed in the present discussion.

comprehensive attitudinal surveys among high school students in Oregon and Alberta made during the past three years.

THE SEVERAL POPULATIONS

Involved in the series of surveys were both Indian and non-Indian students attending an integrated small-town high school in Central Oregon, the identity of which must for the present remain anonymous. Alberta students also involved include: (1) Blood Indian and non-Indian students attending the partially integrated small-town high school at Cardston, Alberta; (2) Blood Indian students attending St. Mary's residential school on the Blood Reserve near Cardston; (3) Blackfoot Indian students attending a partially integrated small-town high school at Gleichen, Alberta; (4) Blackfoot Indian students attending the Chief Crowfoot residential school on the Blackfoot Reserve near Cluny, Alberta; and (5) non-Indian students attending high school in an upper middle class residential area in the city of Calgary, Alberta. For convenience, these several populations have been grouped into five categories: (1) Oregon Indians; (2) Oregon Whites; (3) Alberta Indians; (4) Rural Alberta Whites; and (5) Urban Alberta Whites.

The Oregon data were obtained from populations consisting of 52 Indian and 304 non-Indian students. The Alberta data were obtained from populations consisting of 115 Blood and Blackfoot Indians, 335 rural non-Indian students, and 193 urban non-Indian students. For the most part, these totals represent the usable returns obtained from a questionnaire which was circulated in the classroom among the high school populations in each of the several schools involved.[4] In those instances where the entire student body was surveyed, the net response rate was in no case below 80 per cent.

In terms of basic composition the five populations showed considerable difference in respect to sex, age, class standing, rural-urban residence and father's education.

Differences in sex composition were minimal, with males predominating slightly over females in all instances except the Oregon Indian population in which the proportion of the sexes were approximately six to four in favor of the females. With respect to age, both the Oregon and the Alberta Indian populations showed considerably higher average ages than the three non-Indian populations. In terms of class standing, both Indian populations tended to bulk larger in grades IX and X, with markedly smaller numbers in

4. In two of the five populations, some variation in data collection procedure obtained. Nine of the 115 Alberta Indians were polled at the denominational residence in which they were domiciled while attending a partially integrated public high school in the town of Gleichen. Alternatively, the 193 Urban Alberta non-Indian students were selected primarily on the basis of convenience and availability during a library period on a given day. Consequently neither the degree of numerical nor the social representativeness is fully known.

grades XI and XII, as compared with a much more uniform distribution over the four grades among the two small-town non-Indian populations. Those attending the city high school were drawn from grades X and XI exclusively.

In regard to rural-urban differences in residence, the respective proportions residing on farms were approximately equal, with the exception of the Alberta Indian population in which the relative proportion was twice as large as among the remaining three small-town high school attending populations. Further, only a small proportion of both Indian populations lived in either a town or a city, as compared with more than half of the students in both non-Indian small-town populations. Those attending the city high school were, of course, almost all urban residents.

Respecting father's educational achievement there were again marked differences. The two Indian populations had a much larger proportion of students whose fathers completed only elementary school or less than did the three non-Indian populations. Also, more Indian students gave a "don't know" response on the question pertaining to father's education than did non-Indians.

The Data

It will be convenient to present the data under sub-headings which bear upon the two types of role-taking situations facing high school students. Further, under each of these two sub-headings data will be presented which correspond to the assessed reaction of others and the importance attributed to such reaction by the students themselves. With respect to both types of decisions, the 'significant other' categories which students were asked to evaluate were identical and included the following: (1) parents; (2) school friends; (3) relatives; (4) friends of the family; (5) neighbors; (6) teachers; and (7) other people liked (you like).

Graduation from High School: Two questions were put to the respondents concerning the first of the two role-taking decisions, namely, graduation from high school. The first question referred to the reaction of parents and was phrased as follows: "How disappointed in you would your parents or guardians be if you didn't finish high school?". Answer categories provided were four in number, including "very", "somewhat", "very little", and "not at all". In order to facilitate comparison of the responses to this question with that which follows, those students who reported that their parents would be "very" or "somewhat" disappointed were combined into a single category referred to in Table I as "High Parental Disappointment". Conversely, those students who reported that their parents would evidence "very little" or "no" disappointment are also combined and referred to in Table I as reporting "Little or No Parental Disappointment".

TABLE I

FREQUENCY OF ASSESSED NEGATIVE REACTION ON THE PART OF PARENTS RESPECTING FAILURE TO COMPLETE HIGH SCHOOL TRAINING

	Populations				
	Alberta Indians N=115	Rural Alberta Whites N=335		Oregon Indians N=52	
		Oregon Whites N=304		Urban Alberta Whites N=193	
	A.I.	R.A.W.	O.I.	O.W.	U.A.W.
Response Category	%	%	%	%	%
High Parental Disappointment	90.4	97.0	98.0	99.0	96.3
Little or no Parental Disappointment	9.5	2.9	2.0	1.0	3.6

A.I = Alberta Indians
R.A.W. = Rural Alberta Whites
O.I. = Oregon Indians
O.W. = Oregon Whites
U.A.W. = Urban Alberta Whites

The data reported in Table I show that for each of the five populations the proportions of students reporting high parental disappointment exceeds 90 per cent. Moreover, the range of variation between populations is quite restricted, suggesting that the assessed reaction of parents to the student's potential failure to complete high school is independent of differences in nationality, socio-economic and ecological status.

The second question which was presented to the respondents referred to the reaction of the remaining 'significant other' categories respecting failure to complete high school. The wording of the question was as follows: "Aside from your parents or guardians, would any of these people be disappointed in you if you didn't finish high school?". The response categories provided permitted only a simple "yes" or "no" answer, although in each case a small proportion of the respondents gave no answer to the question. The data reflecting the relative strength of the assessed reaction as measured by frequency of affirmative response of these remaining categories of 'significant others' are shown in Table II.

Table II shows that the students expected far less frequently than was the case with their parents that the remaining categories of 'significant others' would be disappointed at their failure to complete high school. A comparison of Table II with Table I shows the

Table II

CATEGORIES OF REFERENCE RANKED IN THE ORDER OF FREQUENCY OF ASSESSED NEGATIVE REACTION RESPECTING FAILURE TO COMPLETE HIGH SCHOOL TRAINING

Populations

Alberta Indians N=115 Rural Alberta Whites N=335 Oregon Indians N=52
Oregon Whites N=304 Alberta Urban Whites N=193

Reference Category	Percentage Affirmative					Percentage Negative					Percentage No Answer				
	A.I. %	R.A.W. %	O.I. %	O.W. %	A.U.W. %	A.I. %	R.A.W. %	O.I. %	O.W. %	A.U.W. %	A.I. %	R.A.W. %	O.I. %	O.W. %	A.U.W. %
Relatives	80.8	82.3	88.4	95.0	80.3	11.3	12.8	7.6	2.6	16.5	7.8	4.7	3.8	2.3	3.1
Teacher	80.8	78.8	59.6	83.0	54.4	6.0	14.6	17.3	9.0	29.5	13.0	6.5	23.0	8.0	16.0
School Friends	67.8	61.7	63.4	76.0	54.4	20.0	32.2	15.3	16.0	38.3	12.1	5.9	21.1	8.0	7.2
Friends of the Family	49.5	61.7	61.5	79.0	56.4	32.1	29.8	21.1	13.6	32.1	18.2	8.3	17.3	7.3	11.4
Other People You Like	43.4	43.5	61.5	81.6	65.8	31.3	45.0	19.2	11.6	22.2	25.2	11.3	19.2	6.6	11.9
Neighbors	14.7	40.3	28.8	61.6	31.0	61.7	49.5	46.1	28.3	52.3	23.4	10.1	25.0	10.0	16.5

A.I = Alberta Indians
R.A.W. = Rural Alberta Whites
O.I. = Oregon Indians
O.W. = Oregon Whites
U.A.W. = Urban Alberta Whites

proportions reporting that their relatives (the category with the highest frequency of reported disappointment) would be disappointed are in each instance smaller than that reporting parental disappointment. It is noteworthy, however, that among these 'significant other' categories relatives uniformly occupy the position of greatest importance, while neighbors occupy the position of least importance. Even more striking, perhaps, is the median to low position of the category of school friends. Table II shows, further, that the rank order of the categories is identical for Alberta Indians and Rural Alberta Whites, with relatives occupying first position, followed by teachers, school friends, friends of the family, other people liked, and finally, neighbors. In the case of the Oregon Indians the response pattern is roughly the same with the exception of the teacher category which drops from second to fifth place in rank. Among Oregon Whites the response pattern is again similar with the exception of the category "other people liked" which occupies third instead of fifth position in the rank order. Among Urban Alberta Whites the ranking shows considerable variation with "other people you like" ranking in second place, followed by friends of the family in third position, and teachers and school friends being of equal significance and occupying fourth and-or fifth place, respectively.

Turning now to the issue of the importance attributed by students to the assessed reaction of 'significant others', it is unfortunately the case that data are available only for the Urban Alberta White population. The question put to these students was phrased: "Will your decision to finish or not finish high school be influenced to any extent by the fact that the following people might be disappointed in you if you didn't finish?". The categories presented for evaluation omitted parents, but did include school friends, relatives, friends of the family, neighbors, teachers, and other people liked. Again, the response categories called for a "yes" or "no" answer and a small proportion of the students made no answer to the question. These data are presented in Table III.

Reading in the percentage affirmative column of Table III it is apparent that once again relatives are most frequently cited as influential in making a decision concerning completion of high school and neighbors least frequently. The category "other people you like" ranks in second place followed by teachers in third position. School friends and friends of the family occupy fourth and fifth position, respectively. Once again it is perhaps noteworthy that the category school friends occupies a median to low position in the rank order.

TABLE III

CATEGORIES OF REFERENCE RANKED IN THE ORDER OF FREQUENCY OF INFLUENCE RESPECTING THE DECISION TO COMPLETE HIGH SCHOOL TRAINING

	Population		
	Urban Whites N=193		
	Percentage Affirmative	Percentage Negative	Percentage No Answer
Reference Category	U.A.W. %	U.A.W. %	U.A.W. %
Relatives	59.0	36.2	4.6
Other People Liked	39.9	51.3	8.8
Teachers	37.8	51.3	10.8
School Friends	32.6	59.5	7.7
Friends of the Family	30.5	60.6	8.8
Neighbors	9.8	77.7	12.4

U.A.W. = Urban Alberta Whites

Further Training After High School: Two essentially parallel questions were put to the respondents concerning the assessed reaction of others with respect to the role-taking decision to go on to further training after completion of high school. The first question was worded: "If you don't decide to go on to further training (after completing high school), how disappointed in you will your parents or guardians be?". The response categories provided were the same as those used in connection with the question regarding parental disappointment at the prospect of failure to complete high school, and they were treated in Table IV in the same manner as those shown in Table I.

Table IV shows that while students very frequently reported high parental disappointment, the range of variation is somewhat greater than that observed on the parallel set of data reported in Table I. Nevertheless, it will be evident upon examination of Table V below that the frequency of assessed disappointment on the part of parents in each case substantially exceeds that reported by students for other categories of 'significant others'.

Table V shows the data based upon the question concerning assessed reaction of the same list of 'significant other' categories to the students failure to go on to further training after high school. The specific phrasing of the question was: "Aside from your parents or guardians, will any of these people be disappointed in you if you don't go on to further training (after high school)?". Once again, the percentage affirmative column shows that relatives uniformly

occupy the highest rank and neighbors the lowest. Similarily, among Alberta Indians, Rural Alberta Whites, and Oregon Whites teachers occupy second place in the rank order, while among Oregon Indians and Urban Alberta Whites teachers occupy fourth and third position, respectively. The category "school friends" again occupies a median to low position in the rank ordering as do the remaining categories.

Turning again to the issue of the importance attributed by students to the assessed reaction of 'significant others' respecting the decision to go on to further training, there is once again data available only for the Urban Alberta population. In this instance the question put to these students was worded: "Will your decision to go on to further training or not go on be influenced to any extent by the fact that the following people might be disappointed in you if you didn't go on to further training?". The categories of 'significant others' to which the students were asked to make response is identical with those listed in the previous question. The results are given in Table VI.

TABLE IV

FREQUENCY OF ASSESSED NEGATIVE REACTION ON THE PART OF PARENTS RESPECTING FAILURE TO GO ON TO FURTHER TRAINING AFTER HIGH SCHOOL

	Populations				
	Alberta Indians N=115	Rural Alberta Whites N=335	Oregon Indians N=52		
		Oregon Whites N=304		Urban Alberta Whites N=193	
Response Category	A.I. %	R.A.W. %	O.I. %	O.W. %	U.A.W. %
High Parental Disappointment	93.9	94.9	84.0	71.3	90.1
Little or no Parental Disappointment	6.1	5.1	16.0	28.7	9.9

A.I = Alberta Indians
R.A.W. = Rural Alberta Whites
O.I. = Oregon Indians
O.W. = Oregon Whites
U.A.W. = Alberta Urban Whites

Table V

CATEGORIES OF REFERENCE RANKED IN THE ORDER OF FREQUENCY OF ASSESSED NEGATIVE REACTION RESPECTING FAILURE TO GO ON TO FURTHER TRAINING AFTER HIGH SCHOOL

Populations

Alberta Indians N=115 Rural Alberta Whites N=335 Oregon Indians N=52
Oregon Whites N=304 Alberta Urban Whites N=193

Reference Category	Percentage Affirmative					Percentage Negative					Percentage No Answer				
	A.I. %	R.A.W. %	O.I. %	O.W. %	A.U.W. %	A.I. %	R.A.W. %	O.I. %	O.W. %	A.U.W. %	A.I. %	R.A.W. %	O.I. %	O.W. %	A.U.W. %
Relatives	75.6	71.6	67.3	63.3	60.6	16.5	22.6	28.8	33.0	32.6	7.8	5.6	3.8	3.6	6.7
Teachers	71.3	69.5	32.6	62.3	36.7	13.0	22.3	32.6	30.3	49.2	15.6	8.0	34.6	7.3	13.9
School Friends	53.9	37.9	32.6	35.0	21.7	30.4	52.8	40.3	56.6	65.2	15.6	9.2	26.9	8.3	12.9
Friends of Family	49.9	49.2	44.2	38.0	34.7	32.1	40.9	32.6	53.0	51.8	20.8	9.6	23.0	9.0	13.4
Other People You Like	41.7	42.0	42.3	44.3	43.0	34.7	46.2	32.6	46.6	43.5	23.4	11.6	25.0	9.0	13.7
Neighbors	14.7	27.7	11.5	23.3	15.5	53.9	61.4	59.6	65.0	67.3	31.3	10.7	28.8	11.6	17.1

A.I = Alberta Indians
R.A.W. = Rural Alberta Whites
O.I. = Oregon Indians
O.W. = Oregon Whites
A.U.W. = Urban Alberta Whites

TABLE VI

CATEGORIES OF REFERENCE RANKED IN THE ORDER OF FREQUENCY OF INFLUENCE RESPECTING THE DECISION TO GO ON TO FURTHER TRAINING AFTER SCHOOL

Reference Category	Population: Urban Alberta Whites N=193 — Percentage Affirmative U.A.W. %	Percentage Negative U.A.W. %	Percentage No Answer U.A.W. %
Relatives	56.4	36.7	6.7
Other People Liked	33.1	55.9	10.3
Teachers	31.0	57.5	11.4
Friends of the Family	26.9	61.6	11.4
School Friends	21.7	67.8	10.3
Neighbors	8.2	79.2	12.4

U.A.W. = Urban Alberta Whites

Reading in the percentage affirmative column of Table VI, the now familiar pattern of response is once again apparent. Relatives uniformly occupy the position of greatest influence with respect to the decision to go on to further training, while neighbors occupy that of least importance. Again, the category "other people you like" occupies second place, and teachers third place in the rank ordering of the categories. Finally, friends of the family appear in fourth place, while school friends appear in fifth place, being slightly more influential than neighbors.

DISCUSSION

Despite modern rural-urban trends and cultural contact the finding that relatives uniformly occupy the position of greatest importance and neighbors that of lowest in each of the three role-taking situations was somewhat unexpected. Even more surprising, however, was the median to low position occupied by the category of school friends. The recent postulation by Coleman of a more or less independent adolescent culture with values antithetical to the culture of adult society finds little support in the data here reported.[5] Although the ecological conditions which characterize the populations studied no doubt differ somewhat from those studied by Coleman, the present findings suggest that adolescents respond selectively and situationally to the values of the social milieu in

5. James S. Coleman, "The Adolescent Subculture and Academic Achievement", *American Journal of Sociology*, LXV (1960), 337-347.

which they move.[6] The overall uniformity of response pattern among the five different populations in purview, however, lends additional support to this conclusion.

Clearly, the several categories of 'significant others' appear to be taken into account differentially by the students, depending upon the specific nature of the decision facing them. Certain categories of 'significant others' appear to be seen as more concerned with some kinds of role behavior than they are with respect to other kinds. Whether the fact that parents and relatives occupy the position of most extreme assessed reaction and neighbors the least extreme position is a result of differential information or differential identification or both cannot be determined directly from the nature of the data. Similarly, the medium to low position occupied by the category "school friends" raises the question as to whether they have lower visibility, in the sense that there is less informaton avalable to the respondent concerning their disposition, or whether they are adjudged to have dfferent value standards in comparison with alternative categories of 'significant others'. The data, nevertheless, strongly suggest that the students' assessments of the reactions of the several categories of 'significant others' is predicated upon both imputed interest and variegated value standards. Moreover, both are seen as relative to the role-taking issue in purview.

Such an interpretation is warranted by the fluctuation in the relative strength of assessed reaction of the various categories of 'significant others' in the different role-taking situations. Relatives, for example, appear to be regarded as more intensely involved respecting the issue of high school graduation than they are in respect of further training. Correspondingly, the fluctuations in rank order of assessed reaction of the remaining categories of 'significant others' and the degree of importance attributed to such reactions in respect of the different role-taking situations would appear to be accountable on the basis of this same premise.

6. Such an interpretation closely coincides with conclusions reached by Elkin and Westley in their study of adolescent behavior in a Montreal suburb. See Freedrick Elkin and William Westley, "The Myth of Adolescent Culture", *American Sociological Review*, 20 (1955), 680-684.

TOWARDS THE IMPROVEMENT OF METHODS COURSES AND PRACTICE TEACHING

MERRON CHORNY AND ARTHUR G. STOREY

Faculty of Education
University of Alberta, Calgary

A methods course-practice teaching study[1] conducted at the university in Calgary during the 1961-62 academic year failed to disclose any significant relationship between the marks earned in an English methods course and those awarded in practice teaching. The study did, however, discover a positive and significant correlation between changes in student attitudes towards English and skill in teaching. Utilizing the cues provided by this earlier study the present authors set out to improve both the experiences and attitudes afforded a group of students through an English methods course and practice teaching.

The twenty-eight students registered at the university with a teaching major in English were selected as subjects for the experiment. The subjects were divided randomly into two equal groups and assigned to the present authors as student advisees. A committee consisting of the principals of the Elementery and Junior High demonstration schools and the head of the demonstration High School's English Department together with the present authors was organized to delimit the goals to be attempted and outline the procedures designed to achieve them.

In general the aims included[2]:

1. The broadening of practice teaching experiences to include the total demonstration school plant, staff, and student body.
2. The integration of student teachers and teachers with a view to making the students feel they were accepted members of the school staff and the teaching profession.
3. As close a liaison as possible between the English methods course and the experiences provided through practice teaching.
4. A resolution that staff advisors make themselves as available as possible for consultations with their advisees and that their relationships with these advisees be permissive and helpful rather than judgmental.
5. The instigation of seminars involving the members of the committee, demonstration teachers and student teachers following each of the three rounds of practice teaching. The purpose of these seminars was the gathering of opinions in order to obtain feedback for the improvement of both practice teaching and the English methods course.

One of the staff advisors for the experimental group was currently offering the English methods course; the other advisor offered it the previous year. The grades awarded in both the methods course and practice teaching were determined in the same way as for students in general.

Instrument

A forty-five item questionnaire was devised to measure the degree to which each of the above aims was achieved. TABLE I gives the items measuring each aim.

TABLE I
QUESTIONNAIRE ITEMS MEASURING EACH AIM

Aim	Items
1	1- 5
2	6-15
3	16-27
4	28-37
5	38-45

At the conclusion of practice teaching the questionnaire was administered to the experimental group together with a control group matched with them for age, sex, educational standing, marital status, and if married the number and ages of their children.

In part the measuring instrument reads:

Respond to each of the following statements as frankly and as honestly as you can. Do not change any response once you have made it. We want your first and most frank reaction. If a statement is Entirely True of your own practice teaching experience circle the ET category. If it is

Mostly true circle the ..... MT
Somewhat true circle the ..... ST
Neither true nor false circle the ..... U
Somewhat false ..... SF
Mostly false ..... MF
Entirely false ..... EF

Respond to each statement in turn. Do not skip any. Do not put your name on the paper but give us your honest opinion.

STATEMENT	OPINION						
I was able to familiarize myself with the schol's disciplinary methods	ET	MT	ST	U	SF	MF	EF

Procedure

By awarding the U category a weighting of zero and those out from it in either direction weights of 1, 2, and 3 plus or minus depending on the direction both positive and negative scores were obtainable. For example, the scoring for the particular item given above was:

	ET	MT	ST	U	SF	MF	EF	Total
No. of control subjects choosing each category	5	1	10	2	5	5	0	28
Score	+15	+2	+10	0	—5	—10	0	+12
	ET	MT	ST	U	SF	MF	EF	Total
No. of exp. subjects	2	10	8	5	2	1	0	28
Score	+6	+20	+8	0	—2	—2	0	+30

The Chi-square calculated on the hypothesis of equal probability, i.e. the hypothesis that there will be no difference between the total score earned by the two groups, is 7.72, which, with one degree of freedom, is significant above the .01 level. The null hypothesis must therefore be rejected and the conclusion drawn that the experimental and control groups had differential experience insofar

as familiarization with the demonstration schools' disciplinary practices was concerned.

TABLE II summarizes the comparative scores of the two groups on all items.

TABLE II

EXPERIMENTAL AND CONTROL GROUP SCORES

Item	Exp. Gp.	Con. Gp.	D	Chi^2	Sig.
1	43	40	3	.11	Nil
2	45	— 2	47	51.40	.01
3	18	— 6	20	48.00	.01
4	44	37	7	.61	Nil
5	30	12	18	7.60	.01
6	— 1	6	7	7.20	.01
7	9	17	8	2.46	Nil
8	24	37	13*	2.77	Nil
9	30	29	1	.01	Nil
10	35	35	0	.00	Nil
11	40	42	2*	.05	Nil
12	13	5	8	3.56	Nil
13	50	49	1	.01	Nil
14	32	31	1	.01	Nil
15	27	20	7	1.04	Nil
16	41	—31	72	518.40	.01
17	34	—20	54	208.30	.01
18	31	— 9	40	72.73	.01
19	37	—21	58	210.25	.01
20	53	15	38	21.19	.01
21	18	—23	41	336.20	.01
22	23	7	16	8.53	.01
23	44	18	26	10.90	.01
24	73	17	56	34.40	.01
25	60	7	53	41.80	.01
26	32	—17	49	160.07	.01
27	4	13	9*	4.76	.05**
28	45	55	10	1.00	Nil
29	63	57	6	.30	Nil
30	56	42	14	1.92	Nil
31	43	38	5	.31	Nil
32	53	45	8	.65	Nil
33	36	28	8	1.00	Nil
34	49	36	13	1.99	Nil
35	40	36	4	.21	Nil
36	56	53	3	.08	Nil
37	33	35	2*	.06	Nil
38	49	42	7	.54	Nil
39	44	22	22	7.33	.01
40	19	4	15	9.78	.01
41	24	10	14	5.77	.05
42	41	27	14	2.88	Nil
43	34	29	5	.39	Nil
44	47	47	0	.00	Nil
45	41	40	1	.01	Nil
Mean	36.67	20.86			
S.D.	14.80	22.29			

t. of Mean D 3.92 (Sig. .01 level)

* items favoring the control group

** items favoring the control group significantly

TABLE III gives the final marks in practice teaching along with those in English methods.

TABLE III
PRACTICE TEACHING AND ENGLISH METHODS MARKS

Subject	Practice Teaching Mark	English Methods Mark
1	75	65
2	63	61
3	67	62
4	75	70
5	67	41
6	63	47
7	63	47
8	80	63
9	69	65
10	69	60
11	65	65
12	70	80
13	63	60
14	77	74
15	68	68
16	65	56
17	67	61
18	85	71
19	68	81
20	60	51
21	69	66
22	63	84
23	71	55
24	70	72
25	68	55
26	85	80
27	72	73
28	68	57
Mean	69.25	63.79
S.D.	6.59	9.48
Rho coef. of corr.	.54	

Discussion

An examination of TABLE II indicates that aims one and three were reached within the definition of the questionnaire items. Aims two, four and five were not achieved except as sampled by items 6, 39, 40, and 41. It is to be noted that item 27 discriminated between the two groups at the .05 level but in favor of the control group. The t of 3.92 reported in TABLE II indicates that there was an overall enrichment of the experiences offered in both practice teaching and methods course where the experimental group was concerned.

The coefficient of correlation of .54 reported in TABLE III indicates that there was a positive and significant relationship between practice teaching and methods course marks where no significant relationship existed a year earlier.

Recommendations

1. If the study was to be repeated the present items would need to be analysed and altered or discarded as indicated.

2. The Somewhat False and Mostly False categories in the questionnaire are not entirely satisfactory since items in these categories must also be Mostly True and Somewhat True respectively. However, when questioned on this point only one of the fifty-six students responding to the instrument had considered this possibility.

3. More serious is the failure to pre-decide the conditions under which an individual's responses would be considered invalid and discarded. For example, one of the experimental subjects responded to 33 of the 45 items as Entirely False and the other 12 as Entirely True. Both the Chi-square and the Kolmogorv-Smirnov One-sample test indicate the rejection of the hypothesis that such a score could be made by a member of the experimental sample. The elimination of this single paper would have made appreciable differences in the outcome of the experimental findings.

REFERENCES

1. Storey, Arthur G. "Changes in Attitude and Behavior as a Function of Methods Courses," *The Alberta Journal of Educational Research* ,Vol IX, No. 1, March, 1963.
2. Chorny, M. and Storey, A. G., "A Team Approach to Practice Teaching," Unpublished at present.
3. Siegel, M., *Nonparametric Statistics for the Behavioral Sciences,* McGraw-Hill Book Company, Inc., New York, 1956, pg. 47.

AN EXPERIMENTAL APPLICATION OF THE PRINCIPLE OF INCONGRUITY TOLERANCE TO THE COUNSELLING SETTING

G. B. Jones and C. M. Christensen
University of Alberta, Edmonton

Problem

Recent literature in dynamic psychology has involved the postulation of various hypothetical constructs which apparently are related. One group of these concepts centers around the behavioral phenomenon experienced by individuals encountering seemingly opposing cognitive elements (knowledge, beliefs, attitudes or opinions). Such terms as the following have been used by various writers to refer to this phenomenon: ambiguity, inconsistency, uncertainty, conflict, incompatibility, non-conformity, imbalance, asymmetry, dissonance, incongruity and disharmony.

Rokeach (1960) has stated that it would be

> a worthwhile task to examine more closely the extent to which these various concepts overlap each other, the respective ranges of their application, and how they may be more integrated with each other (p. 395).

This "task" was attempted in a fairly extensive review of the literature and because of space limitations will not be reported here other than to note the concepts advanced by a few of the more eminent theorists in the area. From this review, however, an experiment studied the feasibility of placing subjects in a learning situation aimed at promoting "incongruity tolerance". This paper describes the experimental study, its results and its implications for teachers and counselors.

Theoretical Framework

In many cases the aforementioned terms have been used to identify the same concept which describes the existence of conflict in the social, physical or mental environment of the individual. This conflict usually induces tension or emotional arousal causing the individual to take immediate steps to resolve the resultant unpleasant state. It would appear that the following paradigm of incongruity applies to all of the related concepts identified by these diverse terms:

$$\text{incongruity} \rightarrow \text{tension} \rightarrow \text{behavior} \rightarrow \text{congruity}$$

This is a paradigm of operant conditioning; habitual behavior for most individuals.

In this paper then the term "incongruity" refers to the resultant state experienced by the individual when two or more unfitting or dissimilar cognitive elements interact.

The theoretical exposition most fundamental to this study is that of Festinger (1957) who expounded the concept of "cognitive dissonance". Considering pairs of cognitive elements as basic to behavior or knowledge, Festinger considered that dissonance arises when the obverse of one cognitive element follows from the existence of the other element of that pair.

> The existence of dissonance, being psychologically uncomfortable, will motivate the person to try to reduce the dissonance and achieve consonance. (p. 6)

This approach, that the existence of dissonance (or incongruity) will almost automatically lead to a striving for congruity, is similar to that taken by Osgood and Tannenbaum. They use the term "congruity" in relation to the connotational meaning of concepts as measured by Osgood's semantic differential scale. It describes an harmonious unity between an individual's overt attitude to or evaluation of a concept and his judgmental frame of reference. They also use the term to describe the relationship between an individual's evaluative attitude toward a source of information and his attitude toward the concept imparted by this source.

Not too dissimilar but used more in the direction of the act of perception is Bruner's approach (1949). Perceptual incongruity is said to occur when an individual's set or expectancy regarding perception of stimuli is not confirmed. Such incongruity can occur as a result of

> an unexpected concatenation of events, a conspicuous mismatching or an unlikely pairing of cause and effect. (p. 08)

In addition, Taguiri, Bruner and Blake (1958) introduced the term perceptual-affective congruency to describe the extent to which an individual sees the feelings of others for him as corresponding to his feelings for them.

Berlyne (1960) dichotomizes perceptual and conceptual incongruity as two of the aspects of conceptual conflict. Conceptual incongruity, involving symbolic responses, arises when an individual finds that something he has formerly believed authentic and regular now has an exception. Such a concept is not too different from Festinger's "cognitive dissonance". Berlyne defines perceptual conflict, involving perceived stimuli, as arising when stimuli which formerly occasioned certain expectatons on the part of the perceiving organism are now perceived in conjunction with stimuli which disappoint these expectations. A distinct resemblance to statements of Bruner and Postman (1947).

The structure of belief-disbelief system is studied by Rokeach (1960) who reviews this structure in relationship to people's open and closedmindedness. He postulates that the more open one's belief system then the more should incoming information be evaluated on its own merits, the more should the individual be governed by internal self-actualizing forces and the more should he be able to resist externally imposed reinforcement. This definition of openness is considered in this paper as being descriptive of the individual who has approached a level of maximum tolerance of incongruity.

Rokeach also refers to the concept of "belief congruence" noting that

> each person is somehow motivated to arrange the world of ideas, of people, and of authority in harmonious relations with each other. (p. 395)

Further description of maximum tolerance of incongruity is added by Harvey, Hunt and Schroder (1961) who theorize on the structural characteristics of conceptual systems. They have developed a personality theory based on the stage and level of cognitive functioning of the individual. Progression is made from the concretistic or absolutistic stage to the fourth and final stage of development structured with internal and informationally based concepts which have maximum abstractness. An individual at this level of cognitive functioning (as well as one at a level of maximum tolerance of incongruity) would have optimal tolerance of stress or threat, being less likely to resort to defensive resolutions of anxiety than someone at a lower stage. Their statement that the basis for progressive conceptual development to occur is the ability to tolerate refutation of conceptual systems is quite fundamental to the theoretical framework of this present study.

The concept of the interpersonal matrix was introduced by Secord and Backman (1961) to explain personality and behavioral stability and change. This matrix is a composite of three interacting components: an aspect of the individual's self-concept, related interpretations of his own behavior, and his perception of the related aspects of the behavior of the individual with whom he is interacting. Here the individual strives to maintain congruency among these components by shaping the interaction process through manipulation, control or distortion of one or two of the components.

In somewhat a similar approach Heider (1946) spoke of the configuration or unit relationship of person-person, person-nonperson, two persons-nonperson, or three persons as being the fundamental concept in behavioral dynamics. He used the term "symmetrical harmony" to describe the existence of a "balanced configuration" where all the entities have a two way either positive or negative attitude toward each other. These "entities" can include persons,

situations, events, ideas or things. If balance does not exist then forces toward this state will arise and if change is impossible, then tension will result.

It appears that the concepts reviewed do overlap significantly and that the change in terminology by the theorists has paved the way for much repetition of ideas. The existence of congruent states toward which incongruent individuals strive does seem to be accepted as factual by most writers. Complete support also is given to the idea that the existence of incongruity usually produces behavior aimed at its reduction but few writers explained the motivational dynamics of this phenomenon. No mention was made concerning whether this striving is autochthonous and universal or whether it is learned and dependent upon individual differences.

Most of the approaches stressed the noxious aspects of incongruity, apparently evaluating congruous states as valued objectives. Berlyne (1960) and Harvey, Hunt and Schroder (1961) were exceptions. The former pointed out that all aspects of conceptual conflict provide

> the motivation for intrinsic epistemic behavior, and supplement extrinsic motivation when knowledge is sought for practical or social ends. (p. 280)

Such statements indicate the role that can be played by an ability to tolerate incongruity.

Very little mention was made of individual differences in tolerance of incongruity or of the possibility that such tolerance could be profitable for the incongruent individual. Most of the concepts reviewed seemed to regard the process of striving for congruity as a natural outcome of the existence of incongruity: a process initiated by fortuitous circumstances which disappoint the individual's expectations.

Postulates From Related Theory

1. There are individual differences in ability to tolerate incongruity. However, most individuals in Western cultures habitually seek congruous states, desiring to avoid incongruity as easily and as quickly as possible.

2. Both tolerance and intolerance of incongruity are learned personality traits and as such can be changed.

3. If the positive incentives are recognized, an individual can learn to be tolerant of incongruity.

   (a) There will still be an induction of arousal when incongruity is encountered but defensive reactions will not be evoked immediately and perhaps not at all.

(b) An individual, recognizing the advantages of tolerating the incongruity for a period of time, will be able to control and channel the arousal. Progress could be made to a level where control is not necessary and the individual could live with the arousal without tense or restricted feelings. Ultimately the individual could find incongruity quite pleasurable.

(c) Positive incentives for incongruity tolerance could include: realization of a positive type of freedom—"freedom to"—including more inner-directedness (self control) and greater pusuit of personal interests; more openmindedness to novel stimuli and to potentially refuting stimuli; progression toward more abstract stages of conceptual functioning; more creativity and spontaneity; more original and critical thought; and increased epistemic behavior.

Hypothesis

Provision of a learning situation in which the positive incentives of incongruity tolerance are expounded will result in personality changes predictable from the theoretical framework postulated here.

Procedure

Subjects

Forty-one volunteers (twenty-five females) from nine tutorial classes in an introductory course in Educational Psychology at the University of Alberta comprised the sample for this study. The subjects were randomly assigned to one experimental and two control groups. The mean age for the group was 20.6 with a standard deviation of 3.6.

Method

After the first session of testing, two groups met separately for one hour a week for the next seven weeks for informal discussional sessions while the other control group returned only on the ninth week for retesting. The experimental group was led in discussions on the theory behind tolerance of incongruity during the weekly sessions. Attempts were made to have the subjects rephrase the concepts in their own words and provide personal examples of them. Refutation of the concepts provided by the leader was particularly encouraged. Most of each session was assigned to application of the advocated concepts to reality and the subjects were encouraged to think of further personal applications between sessions.

The weekly sessions of the one control group were conducted by the same leader and here discussion centered around topics of personal interest. Most of the discussions stemmed from personal problems or from Educational Psychology course work. All subjects were requested not to discuss group progress with members of another group for the duration of the experiment.

During the retesting session members of the experimental group were also administered an open-ended questionnaire to ascertain their opinions on the topic of incongruity tolerance and intolerance, its personal and educational significance, the effectiveness of presentation of the topic, and personal and interpersonal changes noted as a result of the group discussions.

Testing Instruments

Tests specifically designed to measure personality changes which would result from an increase in incongruity tolerance have not been designed. Therefore a variety of personality instruments was chosen and an attempt was made to compare changes measured by these instruments to those predictable from the theory underlying the concept of congruity.

Rokeach's Dogmatism Scale. This test was constructed to measure individual differences in open and closed belief systems and general authoritarianism and intolerance although no validation information has been obtained regarding its success with the latter two variables. Rokeach states that on the Scale high scorers showed less ability to receive, evaluate and act on relevant information and less ability to form new belief systems. It was felt that subjects who increased in incongruity tolerance would make lower scores on this Scale.

The Dogmatism Scale consists of forty items designed to measure the structure rather than the content of belief systems. Most validity studies have not been too successful. Reliability coefficients ranging from .68 to .93 and clustering around .80 are given for samples of American and English college students, English workers and American veterans.

Edward Personal Preferance Schedule. This instrument provides measures on fifteen personality variables based on two hundred and twenty-five pairs of items. Subjects must indicate the statement in each pair which is most characteristic of their likes and feelings. An attempt has been made to minimize the influence of social desirability on scale values.

Norms (based on normative samples of United States college students) are provided for both sexes for each personality variable and a consistency variable. Quite extensive supportive data has also been accumulated including: coefficients of internal consistency, stability coefficients, intercorrelation of the test's personality variables and validity coefficients gained from comparisons made with the Guilford-Martin Personal Inventory and the Taylor Manifest Anxiety Scale.

Texture Preference Inventory. This test, developed by Dr. Clifford M. Christensen of the University of Alberta ,is composed of one hundred and firty-five pairs of photographs of textures. The method of presentation is similar to that used in Edwards P.P.S. This experimental test, based on the hypothesis that preference for certain types of textures is functoinally related to specific personality variables, gives scores on twelve scales. These scales can be grouped into three types of preferences including: preference for complex, ambiguous or unstructured textures; preference for simple ambiguous textures which could be described as dull and uninteresting; and preference for simple structured textures.

This test is of recent inception and thus, although reliability and validity data have been collected, none has as yet been published. The test was used in a secondary role to compare its results to those attained by the tests of Edwards and Rokeach. It was postulated that subjects who increased in tolerance of incongruity would show an increase in preference for ambiguous textures.

Results

An analysis of variance for a one-way classification, as in Ferguson (1959), was computed to examine the differences between pretreatment and post-treatment scores for all groups on all the variables of the three personality tests. The five and one percent levels of significance were accepted as the desired levels and thus F ratios equal to or greater than 3.25 and 5.21 were required for significance at these levels respectively. The groups are designated by letters as follows: Group "A" is the experimental group, Group "B" is the control group which met in weekly discussional sessions, while Group "C" is the non-attendance control group.

TABLE I

ANALYSIS OF DIFFERENCES BETWEEN PRE- AND POST-TREATMENT SCORES FOR DOGMATISM SCALE

Personality Variable	Group A n=14		Group B n=13		Group C n=14		Anal. of Var'ce. F Test	Signif. Level
	$\bar{d}$	Σd^2	$\bar{d}$	Σd^2	$\bar{d}$	Σd^2		
Dogmatism	1.57	4,390	3.69	4,732	—1.00	3,192	.24	ns

As shown by Table I changes indicated by Rokeach's Dogmatism Scale were far from significant primarily because of large within-group variances.

Table II gives the results compiled on the Edwards Personal Preference Schedule. Significant changes were found for the personal-

ity variables of Affiliation (.005 level), Autonomy (.01 level) and Endurance (.025 level). After treatment the subjects of the experimental Group A scored lower on the Affiliation variable while the control group subjects scored higher over the same period of time. Group A subjects showed greater increases on the Autonomy scale than did the subjects of Group B while those in Group C decreased on this variable. The Endurance variable resulted in no change for Group A subjects while the other groups' subjects attained lower scores on this scale.

Although the findings were not significant, the subjects in Group A did tend to attain lower scores on the Succorance variable and higher scores on the Dominance variable than did the subjects in the control groups.

TABLE II

ANALYSIS OF DIFFERENCES BETWEEN PRE- AND POST-TREATMENT SCORES FOR EDWARDS PERSONAL PREFERENCE SCHEDULE

Personality Variables	Group A n=14		Group B n=13		Group C n=14		Anal. of Var'ce. F Test	Signif. Level
	$\bar{d}$	Σd^2	$\bar{d}$	Σd^2	$\bar{d}$	Σd^2		
Achievement	.79	165	— .15	106	1.0	130	.49	ns
Deference	.07	99	—1.62	167	—1.57	84	1.72	ns
Order	—1.79	225	—1.87	196	— .21	149	.94	ns
Exhibition	2.21	227	.45	112	1.79	183	.93	ns
Autonomy	2.50	245	1.31	117	—1.07	75	5.64	$p < .01$
Affiliation	—2.36	222	2.23	187	1.50	193	7.41	$p < .005$
Intraception	— .57	102	—1.23	218	.64	61	1.33	ns
Succorance	— .93	233	.92	198	0	254	.66	ns
Dominance	2.57	428	.54	161	— .21	119	1.80	ns
Abasement	— .57	268	.51	106	—1.56	172	.69	ns
Nurturance	—2.64	241	0	154	—1.92	233	1.99	ns
Change	.36	109	— .46	136	.86	234	.49	ns
Endurance	.21	119	—3.08	256	— .21	111	4.48	$p < .025$
Heterosexuality	1.28	172	1.92	321	.71	206	.30	ns
Aggression	— .71	148	.92	86	—1.50	159	2.27	ns

TABLE III

ANALYSIS OF DIFFERENCES BETWEEN PRE- AND POST-TREATMENT SCORES FOR TEXTURE PREFERENCE INVENTORY

Variables	Group A n=14		Group B n=13		Group C n=14		Anal. of Var'ce. F Test	Signif. Level
	$\lvert d \rvert$	Σd^2	$\lvert d \rvert$	Σd^2	$\lvert d \rvert$	Σd^2		
I	.29	192	— .23	73	— .71	124	.35	ns
II	2.57	360	1.77	159	1.57	98	.33	ns
III	—1.71	292	1.00	73	—1.57	98	3.16	$p < .06$
IV	—1.07	191	— .15	76	—1.69	123	1.04	ns
V	1.28	90	— .23	71	— .14	128	1.43	ns
VI	2.21	225	— .92	136	1.36	189	3.01	$p < .10$
VII	.07	143	— .46	132	— .21	117	.12	ns
VIII	1.14	172	— .54	83	— .21	127	1.03	ns
IX	—2.29	324	— .54	217	— .43	176	.89	ns
X	—2.36	377	— .77	102	— .07	105	1.46	ns
XI	2.29	340	0	82	—1.14	116	3.63	$p < .05$
XII	—2.00	324	— .51	192	4.14	572	6.77	$p < .005$

NOTE: Variables I, II, V, VIII, XI contain complex ambiguous (unstructured) textures.
Variables III, IV, IX contain simple structured textures.
Variables VII, X, XII contain simple ambiguous (unstructured) textures.
Variable VI contains unstructured textures.

In Table III, significant differences among group scores can be noted on variables XI (.05 level) and XII (.005 level) of the Texture Preference Inventory. After treatment the subjects of Group A tended to increase in preference for the complex-ambiguous textures of variable XI and to decrease in preference for the simple unstructured textures of variable XII. The subjects of Group C showed opposite tendencies while those of Group B remained somewhat the same on both variables.

Once again non-significant trends were noted on a number of the variables of this test including: I, II, V, VI, VIII, IX and X. These changes were in the direction predicted by the basic concepts tested here, for the subjects of Group A consistently tended to

show greater increases in preference for the complex ambiguous textures. They also tended to display a greater decrease in preference for the simple structured and simple abstract textures of the test.

Discussion

It is felt that the findings on the Personal Preference Schedule and the Texture Preference Inventory support the hypothesis that significant personality changes in line with incongruity tolerance theory would result after treatment. The Affiliation variable of Edwards' test purportedly is associated with manifest needs for participation in groups, for formation of friendships and for action in groups rather than individual action. Therefore it is suggested that because the individuals in Group A had learned to tolerate incongruity more adequately and now needed less social support, they scored significantly lower on this variable. These subjects increased significantly on the Autonomy variable and because it is contended that this variable is related to such needs as those for independence in decision making, freedom from convention and conformity, and indulgence in critical thought and action, it is further suggested that additional support is given to the premise that the group increased in tolerance of incongruity.

Further support for the preceding statement was gained from the Texture Preference Inventory. Group A did tend to increase in preference for the ambiguous textures and to decrease in preference for the structured textures. This test was constructed on the assumption that subjects who consistently exhibited preference for clusters of complex-ambiguous or unstructured textures would have fairly high ambiguity tolerance, a willingness to accept a state of affairs which is subject to alternate interpretations or outcomes.

Rokeach's Dogmatism Scale showed no significant changes in personality variables among the groups and either a discrepancy in this present study or in the construct validity of this test could be inferred. It was thought that if the subjects of Group A increased in tolerance of incongruity, they would display more openmindedness. Rokeach implied that persons with open belief systems would score lower on his Scale (and thus be less dogmatic). The results of this study only showed a tendency for the groups which met in weekly discussions to make higher scores after the treatment period. The possibility that many persons could be classified as openminded and that they could support their belief systems quite dogmatically appears to have been eliminated from Rokeach's theoretical framework. It could therefore be suggested that Rokeach's instrument attempts to measure the personality variables of dominance and autonomy and that it does not provide

too reliable a measure of the extent to which individuals are dogmatically closed to new belief systems or the degree to which they are open.

Further experimentation involving this concept of incongruity tolerance would seem necessary. Such experimentation must involve a setting in which the effects of interpersonal, intrapersonal and cultural or normative forces can be studied. An attempt should also be made to use a more representative sample of subjects than was used in the present study. If subjects from cultures other than the Western culture were used it could possibly be determined more positively whether intolerance of incongruity is a universally inherited behavioral trait or one which is acquired from interaction in certain environmental settings. A followup study would also be advisable to determine the retention of incongruity tolerance over a longer period of time. All further experimental attempts would depend upon the development and documentation of an instrument which can provide a valid estimate of the presence and extent of an ability to tolerate incongruity.

The motivational components of incongruity intolerance should also be studied by experimental methods. Perhaps it would be found that the basic motive behind an habitual striving for congruous states is the desire to avoid the feelings of uncertainty and insecurity arising from not knowing or understanding a state of affairs, especially when it is important for the security of the individual involved or when it runs contrary to his expectations.

It could be that tolerance of incongruity is the fundamental principle of all successful psychotherapies and involves nothing more than learning to live with tension, learning to recognize incompatibilities and accept them without defensive reactions. It could be postulated also that incongruity tolerance is a function of intelligence and thus persons at lower levels are not capable of tolerating incongruous states.

The present investigation did not seem to support Rokeach's (1960) apparent combining of the traits of dogmatism and closed-mindedness. Further investigation of this finding could be profitable. Harvey, Hunt and Schroder (1961) seem to disagree with Rokeach also. They claim that the more abstract the conceptual system, the greater the resistance within that system to increasing closedness. They state that the more abstract levels of functioning are characterized by

> the ability to hold strong beliefs and to be ego-involved, while remaining open to alternate evaluations and differentiations, that is, to hold strong beliefs but to be unprejudiced. Such functioning permits ego-involvement without the usual accompanying disadvantages of bias, subjectivity, and distortion. (p. 195)

In educational and psychotherapeutic fields openness of thought and progression toward tolerance of incongruity should be encouraged. In children such progression could be impeded by the encouraging of rigid identification with parental and authoritative models (e.g. teachers), by the absolute presentation of materials and facts with no opportunity allowed for refutation, by requiring that students learn only that which is accepted by the instructor, or by presenting examinations that involve absolute regurgitation of material expounded by the instructor.

Finally it could be implied that incongruity tolerance is positively correlated with creativity, independent and critical thought, and flexibility of personality. It would be further suggested that conceptual structures characterized by such traits are the desirable goals toward which educational and guidance services should encourage students to progress. The basic requirement in providing environmental opportunities for the cultivation of such traits would be that educators and counselors first possess these traits themselves.

REFERENCES

1. Berlyne, D. E., *Conflict, Arousal and Curiosity,* New York, McGraw-Hill Co., 1960.
2. Bruner, J. S., Goodnow, J. J., Austin, G. A., *A Study of Thinking,* New York, John Wiley and Sons Inc., 1956.
3. Bruner J. S., Postman, L., "On the Perception of Incongruity: A Paradigm", *Journal of Personality,* 1949, 18, 206-223.
4. Ferguson, G. A., *Statistical Analysis in Psychology and Education,* New York, McGraw-Hill Co., 1959.
5. Festinger, L., *Cognitive Dissonance,* Evanston, Illinois, Row, Peterson and Co., 1957.
6. Harvey, O. J., Hunt, D. E., Schroder, H. M., *Conceptual Systems and Personality Organization,* New York, John Wiley & Sons, 1961.
7. Heider F., "Attitudes and Cognitive Organization", *The Journal of Psychology,* 1946, 21, 107-112.
8. Heider, F., *The Psychology of Interpersonal Relations,* New York, John Wiley and Sons, Inc., 1958.
9. Osgood, C. E., Tannenbaum, P. H., "The Principle of Congruity in the Prediction of Attitude Change", *Psychological Review,* 1955, 62, 42-55.
10. Rokeach, Milton, *The Open and Closed Mind,* New York, Basic Books, Inc., 1960.
11. Secord, P. F., Backman, C. W., "Personality Theory and the Stability and Change in Individual Behavior, an Interpersonal Approach", *Psychological Review,* 1961, 68, 21-32.
12. Taguiri, R., Bruner, J. S., Blake, R. R., "On the Relation Between Feelings and Perceptions of Feelings Among Members of Small Groups", in *Readings in Social Psychology,* S. E. Maccoby, T. M. Newcomb, E. L. Hartley, New York, Henry Holt & Co., 1958, 110-116.

STEREOTYPING AND THE USE OF ALL-INCLUSIVE TERMS

James G. Snider

University of Alberta, Calgary

The process of stereotyping has been extensively investigated. Lippmann (1922, p. 88) very early pointed out the importance of stereotypes. He said that stereotypes were the pictures we carry in our heads: "We do not so much see this man and that sunset; rather we notice that the thing is man or sunset, and then see chiefly what our mind is already full of on those subjects." And "foreigners," "Baptists," and "Communists" are similar in this sense to men and sunsets. The reviews of Klineberg (1950), Harding, **et al** (1954), and Vinacke (1956) show that stereotyping has many correlates. Stereotyping has been shown to influence ratings of personality traits (Razran, 1950), judgments of political slogans (Lewis 1947), and the identification of group membership from physical appearance (Lund and Berg, 1946). On the other hand, some variables influence the individual's tendency to stereotype. Stautland (1959) found intelligence to be negatively correlated with "strength" of stereotyping. The likelihood of a group being stereotyped may be affected by the actions of that group (Meenes, 1943). Stereotyping may also be affected by the tendency of the individual to use projection as a defense mechanism (Child and Doob, 1943). It is the aim of the present study to delineate still another variable which may influence the individual's tendency to stereotype. The idea for the study arose from the observation that stereotypes are often stated in actual or implied all-inclusive terms. "Schoolboys are lazy"—implying **all** are lazy. "Professors are absent-minded." **All** of them? **Always?** In **every** circumstance? The general question of the study concerned whether or not those who tend toward an excessive use of all-inclusive terms, such as **all, always** and **never,** will not also tend to stereotype more than usual.

Relevance of the Study for Education

It has often been pointed out that it is important in inter-personal relations to guard against stereotyping and important in education to teach against stereotyping (e.g., Klineberg, 1950), and further it has been said that to do these things effectively we need to know as many of the correlates of stereotyping as possible. This study is an attempt to examine for yet one more variable which might help us guard against and teach against stereotyping.

Hypothesis

It was hypothesized that individuals who tend to a more than average use of such all-inclusive terms as **all, never,** and **always,** will have a greater tendency to stereotype than will individuals who tend to a less than average use of such terms. The problem of the study, therefore, was stated as follows: Given two groups, one with a greater tendency than the other to use all-inclusive terms will have a greater tendency to stereotype than will the group with the lesser tendency to use all-inclusive terms.

Design of the Study

In selecting a population for the study, individuals of approximately similar socio-economic condition, educatonal level, and age were sought. Accordingly, the two instruments used were administered to 292 randomly chosen tenth-grade pupils in a large urban high school in Calgary, Alberta.

An instrument devised by the present writer was used as a measure of the tendency to use all-inclusive terms. The test consists of 80 items, 50 "crucial" items and 30 "distractor" items, introduced to insure against pupils discovering the purpose of the test from simply scanning the items, as might be possible, since the items are simple in design. Interviews showed that the distractor items did indeed seem to serve their purpose: none of more than thirty pupils interviewed seemed to have any notion of the true purpose of the test. They were willing to accept at face value that the test was simply to find some "word preferences" of pupils their age.

The "crucial" items of the test were constructed from pairs of words chosen so as to be as nearly alike as possible and yet so that one word of each pair is an all-inclusive term and the other is not and so that the words in each pair were similar in frequency as indicated by Thorndike and Lorge (1944). It is assumed that in such a situation the individual who tends to use all-inclusive terms will tend to prefer the all-inclusive term of each pair. An item-analysis was carried out, utilizing the biserial correlation coefficient, and only items with a coefficient significant at the 5 per cent level of confidence were retained to form the test.

The test has adequate reliability for testing groups: several Kuder-Richardson and split-half reliability coefficients approximate .90. Two validity coefficients were obtained. In a group of 60 ninth-grade pupils, a rank correlation of .47 was obtained between total score and word counts of all-inclusive terms used in free writing. For a group of 48 tenth-grade pupils, the correlation was .36.

The directions for taking the test are simply to choose the preferred word of each pair. The score is the number of all-inclusive terms chosen. The first 20 items of the test, with the all-inclusive terms in the "crucial" items underlined are as follows:

1.	wander	stray	11.	miserable	unhappy
2.	loosen	unfasten	12.	**entirely**	largely
3.	seldom	**never**	13.	**always**	often
4.	detect	reveal	14.	peril	danger
5.	**are**	seem	15.	opinion	knowledge
6.	poverty	desolation	16.	worst	bad
7.	destiny	fate	17.	**must**	can
8.	**certain**	likely	18.	assert	declare
9.	now	soon	19.	maybe	**surely**
10.	den	cave	20.	**completely**	mostly

As a measure of stereotyping, a test devised by Stautland (1959, p. 14) was used. This test, constructed after the manner of Osgood's (1957) semantic differential, places the respondent in the situation of judging various concepts: Englishmen, Chinese, Americans, Mexicans, Russians, Norwegians, in regard to pairs of polar terms: practical-impractical, cruel-kind, intelligent-dumb, inferior-superior, happy-sad, dirty - clean, brave - cowardly, warlike - peaceloving, honest-dishonest, lazy-hardworking. The respondent judges in terms of a seven-point scale, weighted as shown:

Polar Term	(3	(2)	(1)	(0)	(1)	(2)	(3)	Polar Term

Scores on the test may range from 0 to 180, and it is assumed that the higher the score the greater the tendency to stereotype. Stautland obtained test-retest reliability coefficients over .90 with his test and a validity coefficient of .39 with a test of stereotyping devised by Siegel (1954).

In terms of total score on the test of the tendency to use all-inclusive terms, the sample was divided into the upper and lower 25 per cent, giving us two groups, one with a stronger tendency than average to use all-inclusive terms, the other with a lesser tendency than average to use such terms. These two groups were then compared as to their scores on the stereotyping test.

Results

Table I presents a summary of the results of the study. It seems quite clear that the differences shown in the table are not likely to be due to chance alone, as the tests for differences in means indicate. It would appear that the prediction for the study was corroborated. The relationship between the tendency to use all-inclusive terms and stereotyping is apparently a real one and can be empirically demonstrated.

TABLE I

MEANS FOR THE UPPER AND LOWER 25% ON THE TEST OF THE TENDENCY TO USE ALL-INCLUSIVE TERMS AND MEANS FOR SCORES ON THE STEREOTYPING TEST

Test of Tendency to use All-Inclusive Terms	Means		Stereotyping Test Means	
Upper 25%	37.2	$P < .001$	100.7	$P < .01$
Lower 25%	14.4		90.2	

The general tendency of the results shown in the statistical presentation was also observed in the examination of individual cases when pupils had received very high or very low scores on the test of the tendency to use all-inclusive terms. One boy, for example, who obtained the highest score, 49, on the test of the tendency to use all-inclusive terms, also obtained the highest score, 180, on the stereotyping test. The opposite could be seen in pupils who scored very low on the test of the tendency to use all-inclusive terms.

Discussion

While the results of the study suggest that the use of all-inclusive terms is a correlate of stereotyping, it is by no means clear whether or not there is any causal relationship between the two variables. Both the use of all-inclusive terms and the tendency to stereotype may be merely symptoms of a more general "tendency to be all-inclusive." It may be, however, that the use of all-inclussive terms is a contributing factor in stereotyping. More experimental work needs to be done to answer such a question, but the results of the present study suggest that such an undertaking might be worthwhile. Whatever the causal connection between the variables of the study, teachers and others concerned with stereotyping can assume that where there is an excessive use of all-inclusive terms there is also likely to be a tendency to stereotype. If we assume that the use of all-inclusive terms is a contributing factor to stereotyping and that a lessening of the use of all-inclusive terms might curb the tendency to stereotype, we need some means of analyzing the person who uses all-inclusive terms to excess. Fortunately, various therapeutic suggestions (Hayakawa, 1943; Johnson, 1946; Korzybski, 1948), which might be put to use by teachers and others, are available already for dealing with the tendency to an excessive use of all-inclusive terms.

REFERENCES

Child, I. L., and Doob, L. W. Factors Determining National Stereotypes, *J. Soc. Psychol.*, 1943, *17*, 203-219.

Harding, J., Kutner, B., Proshansky, H. and Chein, I. Prejudice and Ethnic Relations. In *Handbook of Social Psychology*, Vol. II, Gardner Lindzey, Ed. Reading, Mass.: Addison-Wesley, 1954.

Hayakawa, S. I. *Language in Action.* New York: Harcourt, Brace, 1943.

Johnson, W. *People in Quandries.* New York: Harpers, 1946

Korzybski, A. *Science and Sanity.* Lakeville, Conn.: The Institute of General Semantics, 1948.

Klineberg, O. Tensions Affecting International Understanding: A Survey of Research, *Soc. Sci. Res. Coun. Bull.*, 1950, No. 62.

Lewis, H. B. An Experiment on the Operation of Prestige Suggestion,in T. M. Newcomb and E. L. Hartley (Eds.) *Readings in Social Psychology.* New York: Holt, 1947, pp. 232-243.

Lippmann, W. *Public Opinion.* New York: Harcourt, Brace, 1922.

Lund, F. H. and Berg, W. C. Identifiability of Nationality Characteristics, *J. Soc. Psychol.*, 1946, *17*, 327-336.

Osgood, C. E., Suci, G. J., and Tannenbaum, P. H. *The Measurement of Meaning.* Urbana, Illinois; U. of Illinois Press, 1957.

Razran, G. Ethnic Dislikes and Stereotypes: A Laboratory Study. *J. Abnorm. Soc. Psychol.*, 1950, *45*, 7-27.

Siegel, S. Certain Determinants and Correlates of Authoritarianism. *Genet. Psychol. Monogr.*, 1954, *49*, 187-229.

Stautland, S. Some Correlates and Determinants of National Stereotyping. Unpublished doctoral dissertation, Stanford University, 1959.

Thorndike, E. L. and Lorge, I. *The Teacher's Word Book of 30,000 Words.* Columbia University Press, 1944.

Vinacke, W. E. *The Psychology of Thinking.* New York: McGraw-Hill, 1952.

Vinacke, W. E. Explorations in the Dynamic Processes of Stereotyping. *J. of Soc. Psychol.*, 1956, *43*, 105-32.

BOOK REVIEW

Team Teaching

Authors: Judson T. Shaplin and Henry F. Olds, Jr. (eds.) and others.

Publishers: Harper and Row, New York, 1964

In the foreword, Frances Keppel asks for a detailed assessment of team teaching and says that "the time has come to sharpen definitions, to puncture some balloons and to put the movement into proper perspective." Most people will agree and this publication makes a commendable effort to accomplish just this.

Written by a number of authors it suffers from some unevenness of style and of subject matter. However, its major strength lies in a clearer definition of team teaching and a development of the theoretical bases on which various aspects of team teaching rely. For example the team concept is related to social organization, small group and organizational theory. Another section developed a useful taxonomy of teams in terms of functional variables.

Educators will be pleased to see also a detailed section on the type of research needed at different stages of team teaching development. Useful information of a different order is given on historical antecedents of team teaching, curriculum adaptation, relationship with the educational reform movement, school buildings and details of many current projects.

Canadian administrators interested in this major innovation of today, as well as the many sitting on the fence waiting, should welcome this appraisal, the main authors of which, though deeply involved in team teaching since its inception six years ago, have allowed their bias to show but rarely.

W. D. Neal

Books Received

The Scottish Scholastic Survey, 1953. University of London Press Limited, 1963. Pp. 211. 25s

A total of approximately 72,400 children ranging in age from 9 years 11 months to 10 years 10 months were given a battery of tests in English and arithmetic. The tests used, keys, norms and variances are presented in detail. A chapter of conclusions and recommendations follows.

Dobinson, C. H., *Schooling, 1963-1970.* George G. Harrap, London and Toronto, 1963. Pp. 175. $3.15.

A critical survey of the British educational system, suggesting changes to meet the needs of changed conditions of life. Discontents are continually expressed with various aspects and levels of the system and with the studies themselves, and suggestions for modification are put forward from many quarters.

Although the author's hard-hitting remarks apply primarily to British education, Canadian educators will find them pertinent to the local scene.

Gardner, D. Bruce, *Development in Early Childhood: The Preschool Years.* Harper and Row, New York. 1964. Pp. 349. $5.75.

A solid introduction to the field, this text organizes the factual information on the child's first six years around a unifying theme of self-development. While recognizing the significance of physical growth, the book gives relatively greater emphasis to achievement of self-hood.

A brief overview prefaces each of the four parts of the text—How We Study Children, Foundations of Development, Aspects of Development in the Preschool Years, and Society of the Pre-School Child.

Gordon, Ira J. *Human Development from Birth through Adolescence.* Harper and Brothers, N.Y. 1962. Pp. 400. $5.75.

This text considers in thorough fashion the development of a newborn infant into an adult. It combines an external approach and an internal approach to the study of the growth of human beings. It will prove useful to those studying and aiding infants in their development into productive adults. The main purpose of the book is to present theoretical and other information which treats the significant relationships among human development in general, general behavior, and the development of the self in each individual.

The Editorial Committee for the Alberta Journal of Educational Research will be pleased to consider research articles from contributors outside the Province of Alberta, for publication in the Journal.

Manuscripts should be addressed to the Editor.

H. E. Smith